Supplementary Information

Identifying hidden nuclear material in orbit from its own neutron spectrum

Yasser Maghrbi

Contents

---

1 Per-neutron spectral separation

For two hypotheses with normalized reconstructed spectra p_A and p_B, let $\ell(E) = \log[p_A(E)/p_B(E)]$ be the per-neutron log-likelihood-ratio score, and let $m_h = \mathbb{E}_h[\ell]$ and $q_h = \mathbb{E}_h[\ell^2]$ be its first and second moments under hypothesis h. The per-neutron spectral separation is

$$\delta^2 = \frac{(m_A - m_B)^2}{q_A + q_B}. \quad (1)$$

For a Poisson inspection of mean exposure λ the accumulated score separates with mean $\lambda(m_A - m_B)$ and variance $\lambda(q_A + q_B)$, so in the Gaussian limit the area under the receiver-operating characteristic is $\Phi(\sqrt{\lambda\delta^2})$ and the exposure required to reach 0.90 is $\lambda_{90} = [\Phi^{-1}(0.90)]^2/\delta^2 = 1.6424/\delta^2$.

δ^2 is a squared separation index, not a Shannon or Kullback–Leibler information. The information decomposition quoted in the main text is a separate quantity and is reported in Section 7.

The Gaussian expression was checked against direct Monte Carlo sampling. For the uranium against count-matched tungsten comparison the two agree to about 1.2 per cent: $\lambda_{90} = 3169$ from direct sampling, 3131 from the Gaussian form.

2 Template statistics and convergence

Spectral templates are histogram estimates of the reconstructed energy distribution under each hypothesis. Since the likelihood is evaluated on held-out realizations, limited template statistics reduce the measured discrimination and make the corresponding exposure conservative. Table S1 gives the exposure required to reach an area under the curve of 0.90 for the baseline configuration against the number of training events per hypothesis.

Table S1: Convergence of the identification exposure with template statistics. Baseline configuration, evaluated on withheld transport realizations. This convergence diagnostic uses twelve bins as a sensitivity configuration; the headline analysis uses eight.

Training events per hypothesis	λ_{90}
700	113
1 400	91
2 800	78
5 600	74
11 000	72
22 000	70
Full sample	70

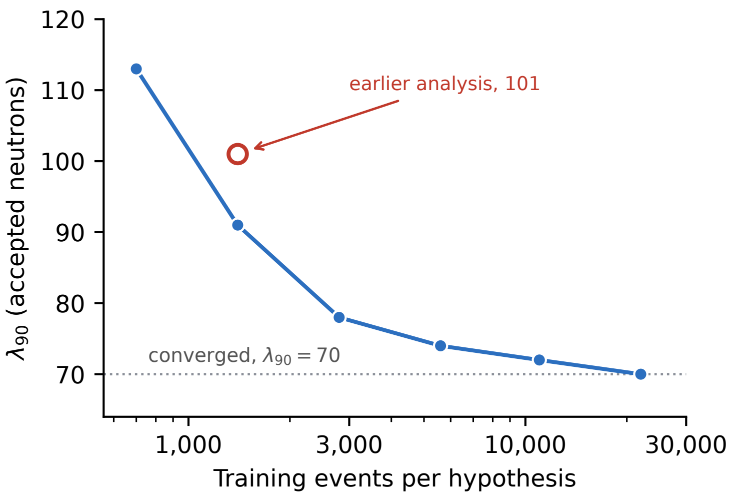

Figure S1: Identification exposure against template size, baseline configuration. The value plateaus above roughly 10^4 training events. The open red marker shows an earlier analysis of the same data at the same template size, using a single limited-template split; it lies within the 16–84 per cent range of the present analysis at that size.

The value plateaus above roughly 10^4 training events, and all exposures quoted in the main text use converged templates. The 1 400-event entry is the median over the present source-disjoint partitions and carries a 16–84 per cent range of about 80–114. An earlier analysis of the same data, using a single limited-template split, returned $\lambda_{90} \approx 101$, which lies inside that range. Both values are conservative: at that template size the measurement is limited by estimation noise, not by the spectral separation.

3 Cross-validation protocol

Six independent transport realizations were generated for each configuration in the high-Z specificity analysis, with five detector realizations per transport realization. The geometry comparison used six realizations for the baseline configuration and four each for the compact and extended configurations, again with five detector realizations per source.

Source realizations were separated first: for every partition of the available realizations into disjoint training and evaluation sets, templates were constructed from the training realizations and the likelihood evaluated on the withheld ones. Detector realizations and their events were pooled only within the source set to which they belonged, so no source spectrum contributed to both sides. For the six-realization configurations all twenty directional three-versus-three partitions were retained. The four-realization compact and extended configurations use all four three-versus-one partitions, with a single held-out realization in each.

Quoted intervals are the 16th–84th percentiles across partitions. Because the partitions share training data they are split-sensitivity bands, not confidence intervals.

Tie correction. All areas under the curve use the tie-corrected mid-rank Mann–Whitney estimator. The count observable is integer-valued and produces many tied scores; a sequential-rank estimator biases tied scores downward and can return values below 0.5 for a classifier carrying genuine information. With the corrected estimator the count classifier in the compact configuration returns exactly 0.500 at every exposure, as required when the two emission rates are equal by construction. The tie-corrected count areas under the curve for the baseline configuration are 0.546, 0.567, 0.597, 0.646, 0.700 and 0.777 at $\lambda = 5$, 10, 20, 50, 100 and 200.

Table S2: Per-neutron spectral separation and the corresponding exposure. Medians over twenty partitions, with 16–84% split-sensitivity bands.

Comparison	δ^2	λ_{90}
Uranium vs conventional structure	0.0235 [0.0227–0.0248]	70 [66–72]
Uranium vs volume-matched tungsten	0.0028 [0.0020–0.0033]	585 [500–840]
Uranium vs count-matched tungsten	0.00052 [0.00035–0.00061]	3169 [2747–4686]

The signed separation was positive in every tested partition for all three comparisons.

4 Geometry dependence

The structural aluminium mass was held fixed and redistributed over three wall thicknesses. Self-shielding weakens both the residual rate difference and the spectral contrast.

Table S3: Identification exposure by structural geometry, converged templates. Standoffs are reference-equivalent values under inverse-square scaling from the four-kilometre detection reference.

Geometry	δ^2	λ_{90}	Standoff
Extended (thin-walled)	0.0439	37	1.5 km
Baseline	0.0235	70	1.1 km
Compact (thick-walled, self-shielded)	0.0091	180	0.7 km

5 Count matching of the high-Z alternative

The count-matched tungsten configuration retains the payload thickness of the uranium reference and enlarges the footprint to 1900 cm^2, corresponding to 183.7 kg of tungsten against 94.5 kg of uranium. Parity was assessed from the full-chain accepted rate, defined as the scored escaping-neutron fluence multiplied by the reconstructed detector acceptance, with identical numbers of incident detector histories for both configurations so that the common normalization cancels.

Table S4: Ratio of tungsten to uranium full-chain accepted rate, evaluated independently for six paired transport realizations.

Realization	R_{count}
1	0.997
2	1.020
3	0.976
4	0.948
5	1.020
6	0.996
Mean	0.993
Seed-to-seed standard deviation	0.027

Thickening the payload at fixed footprint does not reach parity. The accepted rate saturates near 0.81 of the uranium value across a factor of two in thickness, since neutrons escape most readily through the thin dimension and added material increases production and self-absorption along the same path. Widening the footprint at fixed thickness adds material without lengthening that path, and reaches parity at 1900 cm^2.

6 Attribution operating point

Material attribution was defined by two simultaneous requirements: at least ninety per cent probability of correctly identifying the uranium configuration, and at most five per cent probability of attributing uranium to the count-matched tungsten alternative. Thresholds were calibrated on one synthetic inspection sample and evaluated on an independent sample drawn from the same held-out realizations, so the quoted false attribution is measured rather than imposed. Across the tested exposures the achieved rate was 0.049–0.051 against the nominal 0.05, and the uranium power increased monotonically with exposure.

Table S5: Exposure required against the count-matched tungsten alternative at several false-attribution budgets. Medians over twenty partitions.

Criterion	Accepted neutrons
AUC 0.90	3169 [2747–4686]
90% power, 20% false attribution	4501 [3862–6648]
90% power, 10% false attribution	6670 [5458–9604]
90% power, 5% false attribution	8178 [7004–12064]
90% power, 1% false attribution	13173 [11056–19172]

7 Robustness of the spectral statistic

Scalar summaries. A classifier using only the mean reconstructed energy of the accepted neutrons reaches an area under the curve of 0.648 at $\lambda = 100$ in the baseline configuration, against 0.899 for the full-spectrum likelihood under identical sampling and split-sample training. The full distribution retains discrimination not captured by either the count or the mean reconstructed energy.

Rate versus shape. Restoring the Poisson rate term to form the complete joint likelihood ratio changes the area under the curve by at most 0.018 in the baseline configuration, at most 0.003 in the extended configuration, and by exactly zero in the compact configuration, where the two emission rates are equal by construction. The expected log-likelihood ratio of the marked Poisson process separates into a rate term $\mu_A \log(\mu_A/\mu_B) - (\mu_A - \mu_B)$ and a shape term $\mu_A D_{\mathrm{KL}}(p_A \| p_B)$ (main text ref. 21); the shape term carries 92 per cent of the total in the baseline configuration, 99 per cent in the extended and all of it in the compact. Both terms scale with exposure, so the fraction is independent of λ. For the underlying distributions, coarse-graining cannot increase the Kullback–Leibler divergence, so the binned shape fractions are conservative in the population limit.

Number of energy channels. Coarse energy channels recover most but not all of the discrimination. At $\lambda = 100$ in the baseline configuration, two broad channels give 0.829, three give 0.865 and six give 0.884, against 0.903 for the full twelve-bin likelihood. Approaching the full-spectrum result requires of order six energy channels.

Bin number. Repeating the analysis with eight, twelve, sixteen and twenty quantile bins changes the baseline exposure by less than the split-sensitivity band. Bin probabilities were floored at 10^{-6} before renormalization to bound the contribution of sparsely populated bins.

Geometry transfer. Geometry-transfer templates, built without a given structural configuration and applied to an independent transport realization of it, recover comparable discrimination across the tested configurations. In the compact case the transferred template performs slightly better than the matched one, because pooling across configurations reduces the estimation noise of a template built from a weak spectral difference.

8 Structural composition

The count degeneracy is a property of the aluminium structure, not of a minor high-Z constituent. The original bus model used a mixed structural material containing copper alongside aluminium, carbon and hydrogen. Replacing it with pure aluminium at the same structural mass changes the matched-pair accepted count ratio by about one per cent, and the near-equality of the two configurations survives. All results in the main text use the pure-aluminium structure, one of the principal conventional spacecraft constituents identified in the detection proposal (main text ref. 4), together with hydrogen-bearing materials.

9 Robustness to detector calibration

The uranium–tungsten distinction is largely a difference in spectral softness, and an error in the assumed energy scale mimics exactly that. This test therefore asks whether a plausible calibration discrepancy between the simulation and a real instrument could reproduce the signal.

Templates were held at nominal calibration and the evaluation spectra were distorted. This is the realistic failure mode: the template is what the analyst believes, and the data are what the instrument actually delivers. Three distortions were applied,

$$E \to (1+\epsilon)E, \qquad E \to E\left(\frac{E}{10\ \mathrm{MeV}}\right)^{\delta}, \qquad E \to E\left(1+\mathcal{N}(0,\sigma)\right),$$

representing an energy-scale error, an energy-dependent calibration tilt, and additional resolution beyond that assumed in the detector model. All were evaluated over the same twenty cross-validation partitions.

Table S6: Identification exposure under distortion of the evaluation spectra, computed from the Gaussian approximation $\lambda_{90} = 1.6424/\delta^2$; the nominal value of 3131 therefore differs from the direct Monte Carlo threshold of 3169 quoted in the main text. The signed separation remained positive in every case.

Distortion	Magnitude	λ_{90} (U vs structure)	λ_{90} (U vs count-matched W)
Nominal		70	3131
Energy scale	−10%	72	3259
	−5%	71	3379
	−2%	70	3019
	+2%	70	2945
	+5%	70	3217
	+10%	72	3028
Calibration tilt	−0.10	75	4304
	−0.05	72	3491
	+0.05	72	3451
	+0.10	75	4067
Additional smearing	5%	70	3407
	10%	71	3717
	20%	76	4067

A common energy-scale error of up to ±10 per cent changes the count-matched exposure by less than eight per cent. The largest degradation, 37 per cent, comes from a calibration tilt of ±0.10, which displaces the energy axis by about 60 per cent across the analysis range and is far larger than any plausible instrumental error. An additional 20 per cent fractional smearing, which combines in quadrature with the 20 per cent already present in the detector model, costs 30 per cent in exposure. The comparison against conventional structure is almost unaffected throughout.

The signed separation remained positive under every distortion. No calibration error tested reverses the distinction, and none reproduces it from nothing.

This test covers the detector. It does not address variation in the interrogating proton spectrum with orbital position and geomagnetic conditions, object attitude, structural geometry beyond the three configurations examined, or residual background. Those would require additional transport calculations and are not claimed here.

10 Object geometry

Table S7: Dimensions of the simulated configurations. All objects are hollow aluminium cubes of 180 cm external side, centred on the origin, with the payload as a rectangular slab at the centre of the cavity. Neutrons are scored on a spherical surface of 400–401 cm radius.

Configuration	Wall thickness	Payload	Mass
Structure-only, baseline	1.9470 cm	—	1000 kg Al
Nuclear-bearing, baseline	0.66 cm	$31.54 \times 31.54 \times 5.01$ cm	344 kg Al + 94.5 kg U
Structure-only, compact	4.6356 cm	—	1000 kg Al
Nuclear-bearing, compact	1.5215 cm	as baseline	346 kg Al + 94.5 kg U
Structure-only, extended	1.0814 cm	—	1000 kg Al
Nuclear-bearing, extended	0.3720 cm	as baseline	346 kg Al + 94.5 kg U
Volume-matched tungsten	0.66 cm	$31.54 \times 31.54 \times 5.01$ cm	344 kg Al + 96.2 kg W
Count-matched tungsten	0.66 cm	$43.59 \times 43.59 \times 5.01$ cm	344 kg Al + 183.7 kg W

The compact and extended configurations use cube sides of 120 and 240 cm respectively, with wall thicknesses chosen to hold the structural aluminium mass fixed. The uranium payload is defined by atomic number and density alone, so FLUKA assigns natural isotopic composition. Tungsten uses the predefined material at 19.30 g cm^{-3}.

11 Production benchmark

To place the heavy-target ordering in an experimentally constrained regime, thin targets of uranium, tungsten, aluminium and lead of 1 g cm^{-2} areal density were simulated under 800 MeV protons, reproducing the geometry of the published differential (p, xn) measurement (main text ref. 12). Escaping neutrons were scored in four polar bands of $\pm5^\circ$ centred on the measured angles of 30, 60, 120 and 150 degrees, over 0.3–800 MeV, and converted to double-differential cross sections using the band solid angle and target areal density. No parameter was adjusted to the measurement. The simulated uranium target carries natural isotopic composition, whereas the published measurement used depleted uranium. This small isotopic difference is not expected to alter the qualitative heavy-target ordering, but it is not separately tested here.

The uranium-to-tungsten ratio is largest in the 2–8 MeV evaporation region, where it approaches two, and decreases monotonically with neutron energy to approximately 1.3 above 30 MeV. The same energy dependence appears for lead at lower amplitude. A uniform enhancement in total yield would produce an energy-independent ratio; the observed decline with energy shows that uranium's relative enhancement is concentrated at low neutron energy, in the region carrying the orbital spectral discrimination reported in the main text. These calculations do not identify which reaction channel produces the excess. The aluminium target lies substantially below both heavy targets, with the largest contrast at low and intermediate neutron energies.

The few sub-unity ratios occur only in the 150–800 MeV band, where the cross sections

Table S8: Simulated band-integrated neutron production cross sections at 800 MeV, in b sr^{-1}. These are FLUKA results in the measurement geometry, not values extracted from the published data.

Target	Angle (deg)	0.3–2 MeV	2–8 MeV	8–30 MeV	30–150 MeV	150–800 MeV
U	30	0.860	1.296	0.431	0.441	0.337
	60	0.858	1.275	0.391	0.291	0.079
	120	0.834	1.145	0.319	0.098	0.0033
	150	0.811	1.165	0.300	0.083	0.0029
W	30	0.514	0.679	0.285	0.330	0.257
	60	0.493	0.654	0.249	0.226	0.059
	120	0.470	0.592	0.195	0.071	0.0035
	150	0.470	0.568	0.184	0.051	0.0022
Pb	30	0.610	0.842	0.339	0.380	0.279
	60	0.613	0.828	0.281	0.258	0.069
	120	0.600	0.744	0.225	0.080	0.0035
	150	0.567	0.710	0.199	0.058	0.0032
Al	30	0.0113	0.0239	0.0234	0.0433	0.0627
	60	0.0103	0.0222	0.0181	0.0294	0.0105
	120	0.0105	0.0175	0.0116	0.0056	0.00028
	150	0.0099	0.0187	0.0092	0.0041	0.00025

are smallest and the simulated ratios least stable. They are not interpreted as reversals of the heavy-target ordering.

The original measurement (main text ref. 12) reports good agreement between intranuclear-cascade evaporation transport and the uranium, tungsten and lead data over most of the measured range, with significant discrepancies confined to light elements. These results therefore support the production physics relevant to the present comparison, but do not constitute a point-by-point validation of the complete orbital transport and detector chain.

12 Reference-equivalent dwell times

Exposure thresholds were converted into observation times using the published benchmark of five expected accepted neutrons in 7.2 days at four kilometres for a single instrument (main text ref. 4). Counts were taken to accumulate linearly with dwell time, to scale inversely with the square of standoff distance, and to add linearly across identical inspectors, giving $T_c(r, N) = T_0(\lambda_c/\lambda_0)(r/r_0)^2/N$. The instruments are treated as independent inspectors whose accepted events are pooled; no inter-spacecraft coincidence is assumed.

These are reference-equivalent values obtained by scaling a published detector rate. They assume the same source normalization, detector efficiency and live time as that benchmark, and they do not address relative navigation, collision risk, station-keeping or the political acceptability of close approach.

13 Simulation and analysis settings

Table S9: Uranium-to-alternative production ratios, from the band-integrated cross sections of Table S8.

Ratio	Angle (deg)	0.3–2 MeV	2–8 MeV	8–30 MeV	30–150 MeV	150–800 MeV
U/W	30	1.67	1.91	1.51	1.34	1.31
	60	1.74	1.95	1.57	1.29	1.33
	120	1.78	1.93	1.64	1.38	0.95
	150	1.73	2.05	1.63	1.63	1.30
U/Pb	30	1.41	1.54	1.27	1.16	1.21
	60	1.40	1.54	1.39	1.13	1.14
	120	1.39	1.54	1.42	1.22	0.93
	150	1.43	1.64	1.51	1.42	0.92
U/Al	30	76.0	54.3	18.4	10.2	5.4
	60	83.5	57.5	21.6	9.9	7.5
	120	79.1	65.4	27.5	17.5	11.7
	150	82.1	62.3	32.7	20.2	11.3

Table S10: Reference-equivalent dwell times for a constellation of ten instruments.

Claim	λ	at 1 km	at 500 m
Neutron anomaly detected	5	1.1 h	16 min
Distinguished from conventional structure	70	15 h	3.8 h
Distinguished from count-matched high-Z	3169	28.5 d	7.1 d
Uranium-like attribution, 90% power, 5% FA	8178	73.6 d	18.4 d

Table S11: Settings used throughout. Values not listed are the code defaults.

Transport code	FLUKA 4-5.2, `DEFAULTS PRECISIO`
Neutron data below 20 MeV	installed point-wise library, JEFF-3.3
Detector simulation	Geant4 11.4.2 via grasshopper v11.2.2-22-g4368420
Interrogating field	IRENE AP9, tabulated 0.2–2 GeV, peak 0.2 GeV injected isotropically on an enclosing sphere
Primaries per transport realization	2×10^5
Transport realizations	6 (high-Z analysis, baseline geometry) 4 (compact and extended geometries)
Detector realizations	5 per transport realization
Detector histories	10^6 per realization
Detector timing resolution	1.5 ns Gaussian standard deviation, applied to the interplane time difference
Detector spatial resolution	10 mm
Detector energy resolution	20 per cent, fractional
Deposit threshold	0.2 MeV in each plane
Directional acceptance	$\cos\theta_{\text{error}} > 0.95$
Energy bins	8 quantile-spaced (12, 16 and 20 as checks)
Probability floor	10^{-6} before renormalization
AUC estimator	tie-corrected mid-rank Mann–Whitney
Synthetic inspections	1.5×10^4 per exposure, 3×10^4 at operating points
Cross-validation	20 three-versus-three partitions (6 realizations) 4 three-versus-one partitions (4 realizations)